MRC Technical Summary Report #2478

A SMALL SAMPLE EVALUATION OF A BAYESIAN DESIGN METHOD FOR QUANTAL RESPONSE MODELS

Tom Leonard

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

February 1983

(Received November 15, 1982)

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN-MADISON
MATHEMATICS RESEARCH CENTER

A SMALL SAMPLE EVALUATION OF A BAYESIAN DESIGN METHOD FOR QUANTAL RESPONSE MODELS

Tom Leonard

Technical Summary Report #2478
February 1983

ABSTRACT

Leonard proposes a design measure which may be used sequentially to choose the next dose level in a linear logistic quantal response model for bioassay. His design measure averages the posterior distribution of the effective dose over those LD values which are regarded as important. In this evaluation the mode of the design density is used as the next design point, and it is supposed that all LD values between LD 60 and LD 90 are equally important. After ten initial badly designed observations, it is shown that only 20 further, well designed, observations are needed to obtain a design efficiency of about 82%, and an estimated response curve which lies at a maximum of an estimated 6 LD points from the true curve, for all LD values lying between LD 60 and LD 90. If more observations are taken then the design efficiency increases steadily, but it is difficult to increase the accuracy of estimation without either taking many more observations, or by pushing the design points outside the appropriate region. However, within the design region, chosen by any recommended procedure, the method promises excellent robustness, with respect to possible inadequacies in the model, whilst outlying design points would not provide such robustness.

AMS (MOS) Subject Classifications: 62K99, 62F15

Key Words: Quantal response, Bioassay, Effect dose, Design measure, Posterior mean, Bayesian extinction

Work Unit Number 4 (Statistics and Probability)

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

SIGNIFICANCE AND EXPLANATION

Leonard (1982) proposes a design measure for the quantal response model, based upon a mixture of posterior distributions of the effective doses. In this paper the mode of this distribution is evaluated sequentially according to the criteria of "estimation accuracy" and "design efficiency" which relate to both good estimation conditional on the model, and good robustness with respect to deviations from the model. For sequential design points, for ED values between ED 60 and ED 90, it is shown by simulations that, after an initial batch of ten badly chosen design points, only 20 further observations are needed to achieve a design efficiency of 82% and an estimated response curve which lies at a maximum of an estimated 6 ED points from the true curve. For more than 30 observations it is necessary to spread out the design points more, in order to sacrifice some design efficiency for more estimation accuracy. Further simulati_ns are currently being carried out by Leonard and Hamada, to investigate this aspect more closely.

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

A SMALL SAMPLE EVALUATION OF A BAYESIAN DESIGN METHOD FOR QUANTAL RESPONSE MODELS

Tom Leonard

1. INTRODUCTION

Consider zero-case responses $y_1,\ldots,y_n$ which are independent binary variables, with

$$p(y_i = 1|\theta_i) = \theta_i \qquad (i = 1,\ldots,n) . \tag{1.1}$$

Assume the linear logistic model

$$\alpha_i = \log \theta_i - \log(1 - \theta_i) = \beta_0 + \beta_1 x_i \tag{1.2}$$

where $x_1,\ldots,x_i$ will be referred to as "dose levels".

We address the following three problems:

(i) Estimation of the response curve

$$\theta(x) = e^{\beta_0+\beta_1 x}/(1 + e^{\beta_0+\beta_1 x}) \tag{1.3}$$

after n observations.

(ii) Choice of the next design point, x_{n+1}, given $y_1,\ldots,y_n$ and $x_1,\ldots,x_n$, and the sequential choice of x_{n+r}, given $y_1,\ldots,y_{n+r-1}$ and $x_1,\ldots,x_{n+r-1}$, for $r = 1,2,\ldots$.

(iii) The problem in (ii) but where the results, for values of θ in (1.3) lying within specified limits (a,b), are robust against possible inadequacies of the model in (1.2).

These problems may be tackled as follows:

(i) The response curve in (1.3) will be estimated by the approximation recommended by Leonard (1982) to the posterior mean value function of

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

the curve. This mean value function is vastly superior in terms of accuracy of estimation to the maximum likelihood estimate of the response curve, and is primarily responsible for our good practical results. It is also undoubtedly superior to non-parametric estimates based upon Dirichlet processes, since the latter are not smooth enough to reasonably approximate a smooth true response curve, and also depend upon a number of complicated prior parameter values.

(ii) For illustrative purposes it is supposed that the function in (1.3) is of interest for values of θ lying between 0.6 and 0.9. At any stage in the experiment, the next design point will be chosen to be the mode of the design density recommended by Leonard (1982), given all observations up to that point. This density averages the posterior densities of the effective closes over LD values 60 to 90. It may be adjusted to any θ region of interest.

(iii) The design procedure in (ii) attempts to choose x values such that the corresponding $\theta(x)$ values have the best chance of lying between 0.6 and 0.9. Note that

(A) Conditional on the model better estimates of this region of the response curve may be obtained by taking x points further apart.

However,

(B) Results based upon widely spread x points will not be robust under inadequacies in the model. Therefore it is sensible to define design percentage efficiency as the long run percentage of x values such that for the true curve θ, the corresponding $\theta(x)$ values lie between 0.6 and 0.9. A high percentage design efficiency means good estimation and excellent robustness, whilst a low efficiency could mean excellent estimation conditional on the model, but bad robustness.

For an estimate θ^* of the true response curve θ, the <u>estimation accuracy</u> is defined to be

$$\text{DIFFMAX} = \max_{0.6 \leq \theta(x) \leq 0.9} |\theta^*(x) - \theta(x)| \, . \qquad (1.4)$$

The clear objective should be to obtain good estimation accuracy whilst at the same time maintaining high percentage design efficiency. There is a trade-off between estimation accuracy and design efficiency; the objective is to obtain good estimates of θ which are also robust against changes in the model.

2. A SMALL SAMPLE EVALUATION, AND OVERALL CONCLUSIONS

The method described in Section 1 was evaluated based upon a true response curve taking the linear logistic form in (1.3) with $\beta_0 = -4.39$ and $\beta_1 = 8.79$. Owing to properties of the curve, under location and scale transformations, the choices of β_0 and β_1 can be made arbitrarily. The particular choices made ensure that $\theta(0.5) = 0.5$ and $\theta(0.9) = 0.75$.

In each simulation the ten initial x's were taken to be $0, 0.05, 0.1, 0.15, \ldots, 0.45$ corresponding to θ values ranging between 0.012 and 0.392. These points are badly designed by intention in order to give worst-case results. Our estimation accuracies and design efficiencies could be noticeably improved by basing the first few choices on sensible reasoning relative to the practical situation under consideration.

The estimation accuracies for 10 runs, each with 30 further observations based upon well (sequentially) designed x points, are described in Table 1.

Table 1: Estimation Accuracies for up to 30 Further Observations

Sample Size	10	15	20	25	30
Run 1	0.05	0.05	0.04	0.03	0.02
Run 2	0.15	0.09	0.07	0.02	0.02
Run 3	0.12	0.08	0.06	0.11	0.06
Run 4	0.11	0.13	0.10	0.09	0.12
Run 5	0.06	0.04	0.02	0.04	0.09
Run 6	0.03	0.01	0.03	0.03	0.02
Run 7	0.15	0.06	0.04	0.03	0.06
Run 8	0.16	0.06	0.04	0.02	0.03
Run 9	0.08	0.08	0.05	0.05	0.09
Run 10	0.06	0.11	0.11	0.10	0.10
Overall	0.090	0.071	0.056	0.052	0.061

We see that 20, well designed, observations enable us to estimate the true percentage response curve $100\ \theta(x)$ to within under an estimated 6%, of the true curve, for all values of x such that $60\% < 100\ \theta(x) < 90\%$. This estimated accuracy is remarkably low, and is primarily caused by the particular estimation technique employed, which sensibly compensates for large standard deviations of the maximum likelihood estimates for β_0 and β_1.

Beyond 20, well designed, observations it is difficult to substantially improve upon estimation accuracy (together with good design efficiency) without taking a large number of further observations. For example, runs 3 and 4 were continued for a further 30 observations yielding

Table 2: Estimation Accuracies for up to 50 Further Observations

Sample Size	25	30	35	40	45	50
Run 3	0.11	0.06	0.05	0.08	0.08	0.09
Run 4	0.09	0.12	0.08	0.08	0.08	0.09

Furthermore, for all ten runs, the estimation accuracies do not noticeably increase after 25 or 30 observations. See the last two columns of Table 1.

Our tentative recommendations for the practical application of this procedure are:

(A) Choose up to ten initial observations as sensibly as possible, by reference to the practical situation under consideration.

(B) Obtain 20-25 further, well-designed, observations based on the Bayesian procedure.

(C) Then stop, unless it is intended to proceed to a sample size of about 100. (In this case, several e.g. 5 new design points should be chosen at each time stage. These should be spread out according to the percentiles of the design density.)

The percentage design efficiencies for the ten runs are described in Table 3.

Table 3: Percentage Design Efficiencies

Sample Size	10	15	20	25	30
Run 1	90.0	93.3	85.0	96.0	96.7
Run 2	40.0	60.0	70.0	76.0	80.0
Run 3	60.0	73.3	80.0	84.0	86.7
Run 4	40.0	53.3	65.0	72.0	76.3
Run 5	70.0	80.0	85.0	88.0	90.0
Run 6	100.0	100.0	100.0	100.0	100.0
Run 7	60.0	66.7	75.0	80.0	83.3
Run 8	40.0	60.0	70.0	76.0	80.0
Run 9	90.0	86.7	90.0	92.0	93.3
Run 10	90.0	93.3	90.0	92.0	93.3
Overall	68.0%	76.7%	82.0%	85.6%	91.3%

We see that after 20 well-designed observations, 82.0% of the design points correspond to $0.6 \leq \theta(x) \leq 0.9$ for the true curve. This excellent efficiency increases to 91.3% after 30 observations.

Therefore, as well as getting good estimation accuracy for small sample sizes, the results are based upon design points lying in the important region of the response curve. This promises excellent robustness under changes in the sampling model, in particular when these changes are outside the region of interest.

3. SEQUENTIAL BEHAVIOUR OF DESIGN POINTS

Consider the sequential behaviour of the design points for Run 1. Here the ten badly chosen design points, discussed at the beginning of Section 2, yielded the, rather uninformative, responses, 0,0,0,0,0,0,1,0,0,0.

The next 30 design points are described in Table 4.

Table 4: Design Points for Run 1

Designed Observation	Design point	True Response Curve $\theta(x)$	Simulated Response y
1	0.84	0.95	1
2	0.70	0.85	1
3	0.64	0.77	1
4	0.60	0.71	0
5	0.70	0.85	1
6	0.66	0.80	1
7	0.64	0.77	0
8	0.70	0.85	1
9	0.68	0.83	1
10	0.68	0.83	1
11	0.66	0.80	1
12	0.64	0.77	1
13	0.62	0.74	1
14	0.62	0.74	1
15	0.62	0.74	1
16	0.60	0.71	1
17	0.60	0.71	0
18	0.62	0.74	1
19	0.60	0.71	1
20	0.60	0.71	1
21	0.60	0.71	0
22	0.62	0.74	0
23	0.64	0.77	1
24	0.62	0.74	0
25	0.64	0.77	1
26	0.64	0.77	1
27	0.62	0.77	1
28	0.62	0.77	1
29	0.62	0.77	1
30	0.62	0.77	1

The maximum of the badly chosen design points is 0.39, but the Bayesian procedure moves the next point as far away as 0.84. Then observed "ones" tend to push the next design point downwards, whilst observed "zeros" tend to push the design upwards. Zeros create more movement than ones; since these are required to occur between proportions of 0.1 and 0.4 of the time, so that their less likely occurrence has greater effect. The corresponding values of the true response curve move very quickly to the region (0.6, 0.9) and then stabilize near the middle of this region. The downwards movement of the design point, after a positive observation, moves the design into a region where the experiment is likely to be more informative. The general behaviour of the design points seems to be highly sensible.

4. THE POSTERIOR MEAN EFFECT

Note that after the 30 designed trials, Run 1 yielded maximum likelihood estimates of -5.58 and 10.90, for β_0 and β_1, with approximate standard deviations of 2.15 and 4.17. This result is typical of all ten runs, and suggests large discrepancies from the true values $\beta_0 = -4.39$ and 8.79. Nevertheless, the posterior mean of the response curve was in close agreement (estimation accuracy = 0.02) with the true curve, since this adjusts the maximum likelihood estimate of the response curve to allow for the uncertainty represented by the standard deviations.

This dramatic improvement when compared with classical procedures should perhaps be termed "the posterior mean effect."

5. COMPUTER SOFTWARE

Program BIOMAIN.FOR on the VAX system at MRC applies computer simulations (subroutine PGMGEN.FOR) to the design and estimation procedures (subroutine

PGMB10.FOR) when the design region is between 60% and 90%. Program B10.FOR completes the design and estimation procedures when either (i) the design region is unrestricted, or (ii) the design region comprises any single point lying between 1% and 99%.

6. ACKNOWLEDGEMENTS

The author wishes to thank George E. P. Box and J. Richard Moore for useful comments, and Michael Hamada for helping with the computing.

REFERENCE

Leonard, T. (1982), An inferential approach to the bioassay design problem, MRC Technical Summary Report #2416.

TL:scr

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 2478	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A SMALL SAMPLE EVALUATION OF A BAYESIAN DESIGN METHOD FOR QUANTAL RESPONSE MODELS		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Tom Leonard		8. CONTRACT OR GRANT NUMBER(s) DAAC29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 4 - Statistics and Probability
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P.O. Box 12211 Research Triangle Park, North Carolina 27709		12. REPORT DATE February 1983 13. NUMBER OF PAGES 9
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Quantal response
Bioassay
Effect dose
Design measure
Posterior mean
Bayesian extinction

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Leonard (1982) proposes a design measure which may be used sequentially to choose the next dose level in a linear logistic quantal response model for bioassay. His design measure averages the posterior distribution of the effective dose over those LD values which are regarded as impo. tant. In this evaluation the mode of the design density is used as the next design

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 35 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT - cont'd.

point, and it is supposed that all LD values between LD 60 and LD 90 are equally important. After ten initial badly designed observations, it is shown that only 20 further, well designed, observations are needed to obtain a design efficiency of about 82%, and an estimated response curve which lies at a maximum of an estimated 6 LD points from the true curve, for all LD values lying between LD 60 and LD 90. If more observations are taken then the design efficiency increases steadily, but it is difficult to increase the accuracy of estimation without either taking many more observations, or by pushing the design points outside the appropriate region. However, within the design region, chosen by any recommended procedure, the method promises excellent robustness, with respect to possible inadequacies in the model, whilst outlying design points would not provide such robustness.

END

DATE
FILMED

6-83

DTIC